

AD-A275 877



INTRODUCTION PAGE

Form Approved
OMB No. 0704-0186

nated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Drive, Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

DATA PROCESSING CENTER ECONOMICS

Using industry benchmark data to estimate cost savings from efficiency improvements and consolidation.

Competitive Benchmarking is a Total Quality Management technique used by industry to set performance goals. Competitive benchmarking is the quantitative comparison of operations conducted at a DPI independent of the workload being processed. Several firms offer competitive benchmarking of DPIs as a service. These firms maintain models and data bases of DPIs for cost and performance elements that force normalization between installations. The cost and performance of DPI functions -- Central Processing Unit, disk storage, help desk, etc. -- are calibrated. By comparing the values attained by the DPI being benchmarked against industry practices, the DPI can focus the efforts of TQM Process Action Teams (PAT).

The economics of DPIs can be extracted from the comparative data supplied by the benchmark vendors. The vendor's reports have been used by the Defense Information Systems Agency (DISA) to estimate cost savings achievable through efficiency improvements and consolidation.

Dr. James C. Criner
DISA/JIEO/CIM/XII
701 South Courthouse Road
Arlington, VA 22204-2199
(703) 285-5323

Accesion For	
NTIS	CRA&I <input checked="" type="checkbox"/>
DTIC	TAB <input type="checkbox"/>
Unannounced <input type="checkbox"/>	
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and / or Special
A-1	

3

USING INDUSTRY BENCHMARK DATA TO ANALYZE DOD DATA PROCESSING INSTALLATION ECONOMIES OF SCALE

James C. Criner and Dean Halstead

1. Introduction

The Department of Defense's (DoD) corporate Information Management (cIM) initiative has many aspects. Part of the Defense Information Systems Agency's (DISA) support to cIM is demonstration of techniques which bring about superior Information Technology performance. As part of this effort, DISA conducted a Pilot Project to demonstrate the utility of competitive benchmarking of DoD Data Processing Installations (DPI) against commercial DPIs. Four vendors were used in the Pilot Project: Real Decisions Corporation; Nolan, Norton & Co.; Compass America; and KPMG Peat Marwick (Advanced Technology).

Competitive benchmarking differs from acquisition support or workload benchmarking familiar to Government DPI managers. Workload benchmarking is used to determine which hardware configuration best meets a fixed data processing load. Competitive benchmarking, or simply benchmarking, looks for best practice regardless of the nature of the work being done. Benchmarking is the continuous process of comparing one's performance against that of others to determine the best practice and to establish and validate process goals for process improvement. Students of Total Quality Management (TQM) will quickly note the use of key TQM concepts such as continuous process improvement. Successful benchmarking efforts are generally imbedded within a TQM culture.

Benchmarking brings reality to the TQM process. Resources will be wasted if an organization attempts the impossible. Benchmarked standards and objectives have been achieved by others and hence are known to be feasible. The benchmarks provide the measurements for management by fact, not myth.

Benchmarking helps develop a strategy for improvement. If performance is near best practice, then incremental change -- the customary TQM practice -- is appropriate. Otherwise, radical action, such as Business Process Redesign, is required. The benchmark shows how close performance is to optimum.

Benchmarking enables the Process Action Teams. Benchmarking not only provides the checkpointing needed for implementation, it also provides a basis for performance evaluation. Hence it is an enabler for TQM.

94 2 10 18.8¹

94-04864

27P

The remainder of this paper is separated into two parts. In the first part, we discuss the results of the Pilot Project and describe the efficiency of DoD DPIs when viewed from the perspective of the benchmarking vendors' data bases. We also show how it is important to correctly interpret the results of a benchmarking analysis to ensure that you receive maximum benefit from the results.

In the second part, we further analyze the benchmarking process and explore the relationship between benchmarking and DPI economics. First, we examine the difficulty in determining the size of a DPI -- an essential step in the analysis of economies of scale. A simple measure of computing capacity, such as MIPS (Millions of Instructions per Second) is determined to be an adequate metric. Second is an examination of the impact of the perspective taken by the different benchmarking vendors. The four vendors take distinct viewpoints of a DPI. These differences are examined and the determination made that they are differently useful. Next is a discussion of how overall performance of a DPI is measured. Fourth, is a demonstration of how DPI economics can be extracted from the comparison data provided by the vendor. This comparison data consists of information collected in previous engagements and is used by the benchmarkers to compare a DPI against other DPIs. Finally, we conclude with a discussion of the dramatic 20% per year improvement achieved by Xerox using TQM supported by an aggressive program of benchmarking.

2. Part 1: DoD DPI Benchmarking Results

One of the vendors in the Pilot Project, Peat Marwick, analyzed five DoD DPI ranging in size from 59 to 133 installed MIPS. The primary platform for all five DPIs was IBM/IBM compatible mainframe computers using the MVS operating system. Figure 1 provides some key economic indicators for each site and shows the variation in efficiency that occurs in DoD. The average total cost per used MIPS of \$391,458 is approximately 33% higher than population average, and approximately 137% higher than Peat Marwick's best practice.

Figure 2 shows that one of the explanations for lower DoD efficiency is that DoD tends to not use its computers as heavily as the other DPIs in Peat Marwick's data base. This lower utilization occurs in both the prime and non-prime shifts. DoD does a good job of managing and using its disk space, as shown in Figure 3. However, the total amount of disk space per DPI tends to be larger than the average in the data base. DoD also does a good job of managing tape, as shown in Figure 4. However, several of the DoD DPIs mounted a significantly more tapes when compared to the data base average. Also, tape operator productivity is lower in the DoD DPIs.

DoD DPI's production job failure rate is higher than Peat Marwick's data base average, as shown in Figure 5. However, an interesting fact is that approximately 44.7% of the failures are the Central Design Activity's responsibility. Figure 6 shows that DoD DPIs print about the same volume as DPIs of the same size; however, print operator productivity is lower than the population average.

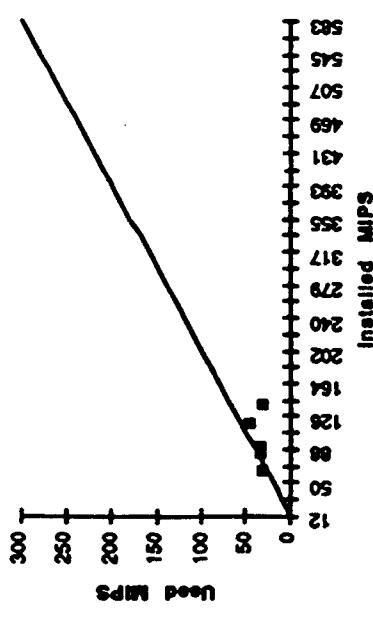
Key Indicators	A	B	C	D	E	Total
Installed MIPS	59	78	116	84	133	470
Overall Utilization %	38.8%	26.1%	31.1%	23.4%	14.2%	26.7%
Prime Shift Utilization %	47.7%	45.8%	45.6%	38.2%	26.8%	40.8%
Used MIPS	23	20	50	20	19	132
Staff	78.3	79.1	128.6	70.4	145.4	501.8
Staff Per Used MIPS	3.44	3.89	3.56	3.60	7.69	4.44
Annual Operating Cost	\$6,910,000	\$6,069,000	\$11,816,000	\$7,448,000	\$12,215,000	\$44,458,000
Cost Per Used MIPS	\$303,892	\$298,505	\$327,290	\$381,014	\$646,589	\$391,458

©Copyright 1992 by KPMG Peat Marwick. All rights reserved.

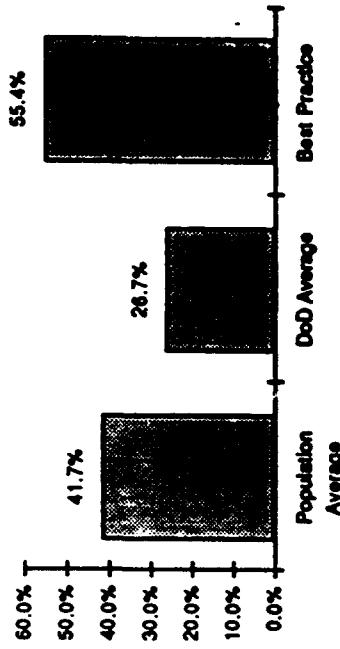
KPMG Peat Marwick
Management Consultants

Figure 1: DoD Key Economic Indicators

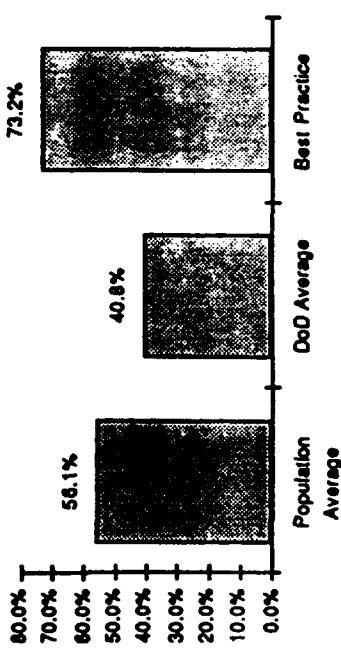
Installed vs. Used MIPS



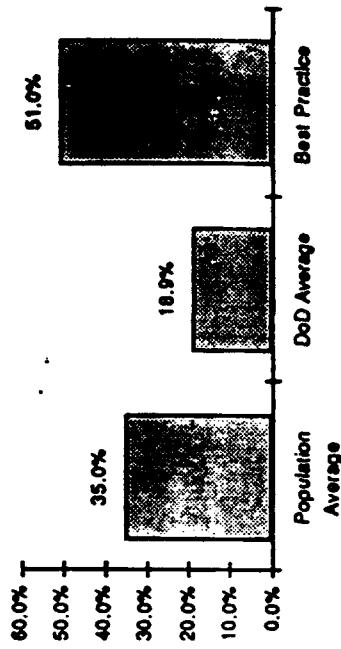
Overall % Used



Prime Shift % Used



Non-Prime Shift % Used



©Copyright 1992 by KPMG Peat Marwick. All rights reserved.

KPMG Peat Marwick

Figure 2: DoD CPU Utilization

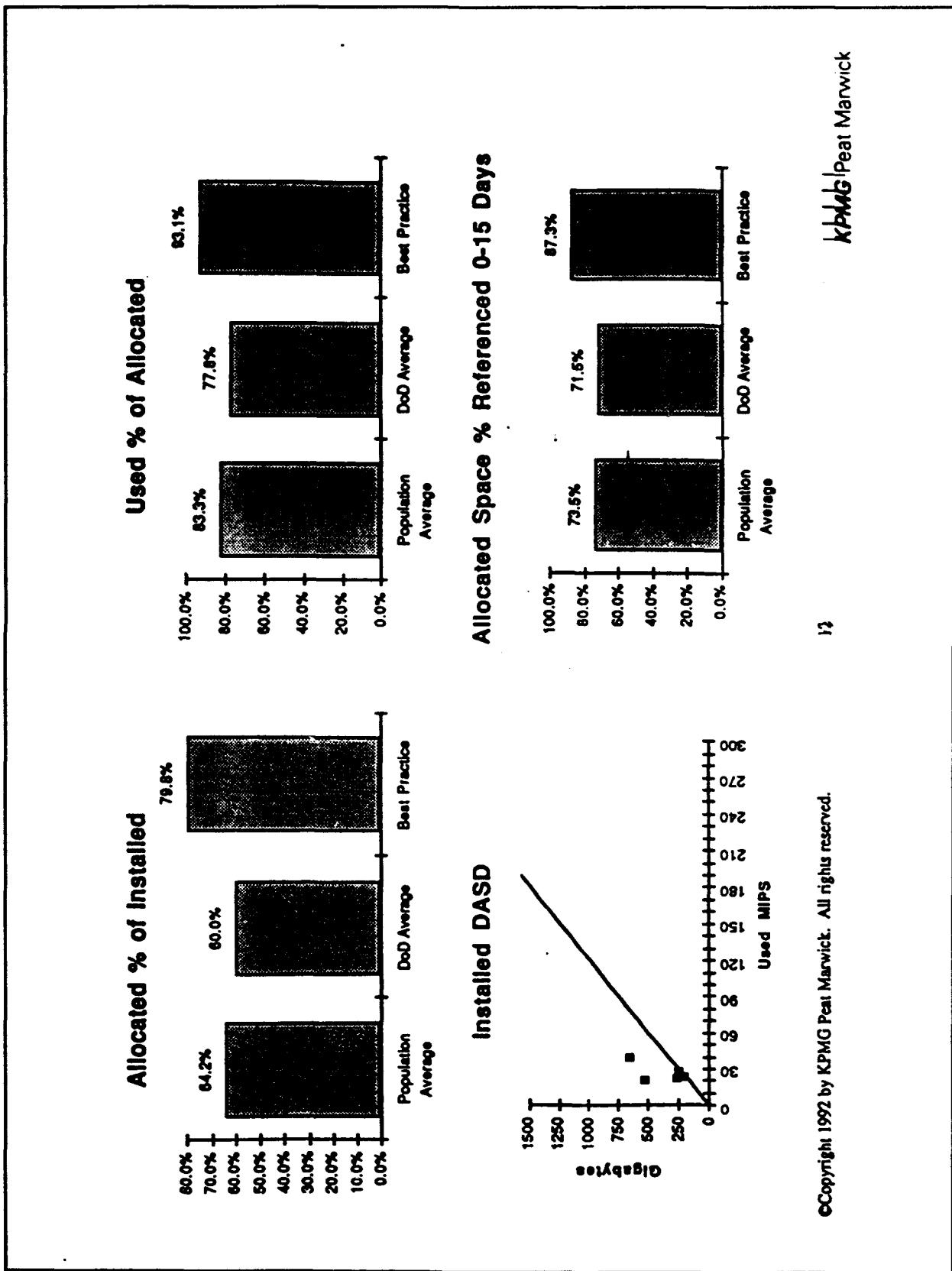
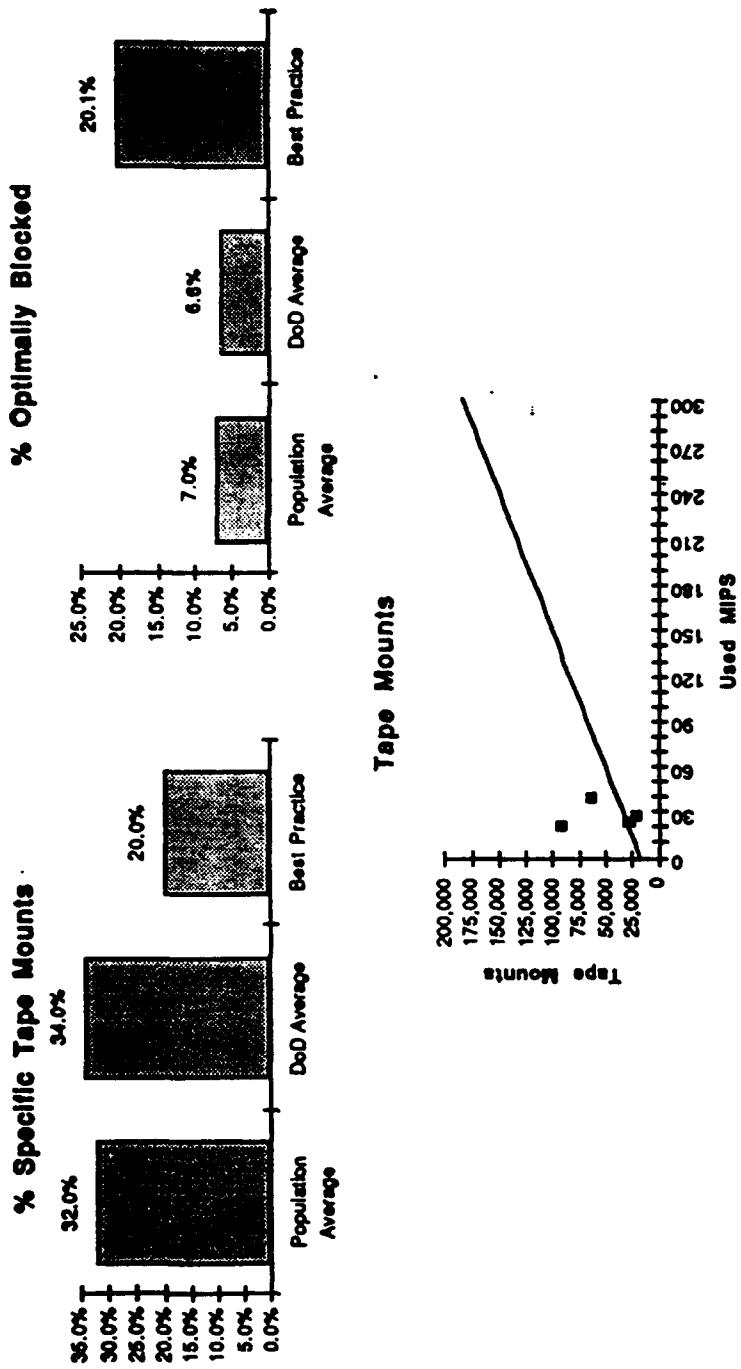


Figure 3: DoD DASD Management



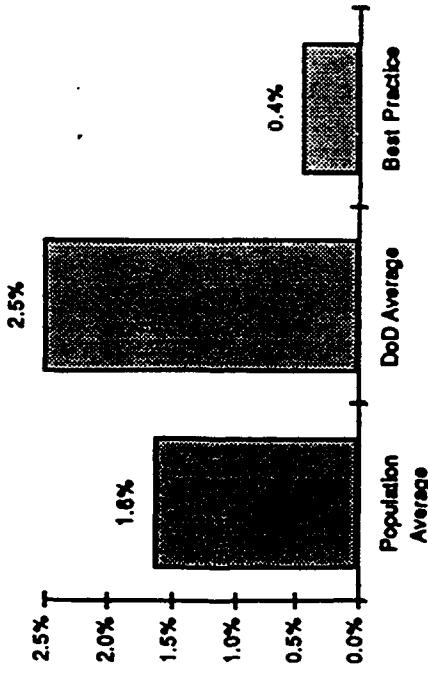
©Copyright 1992 by KPMG Peat Marwick. All rights reserved.

KPMG Peat Marwick
Management Consultants

Figure 4: DoD DPI Tape Management

Reasons for Failures

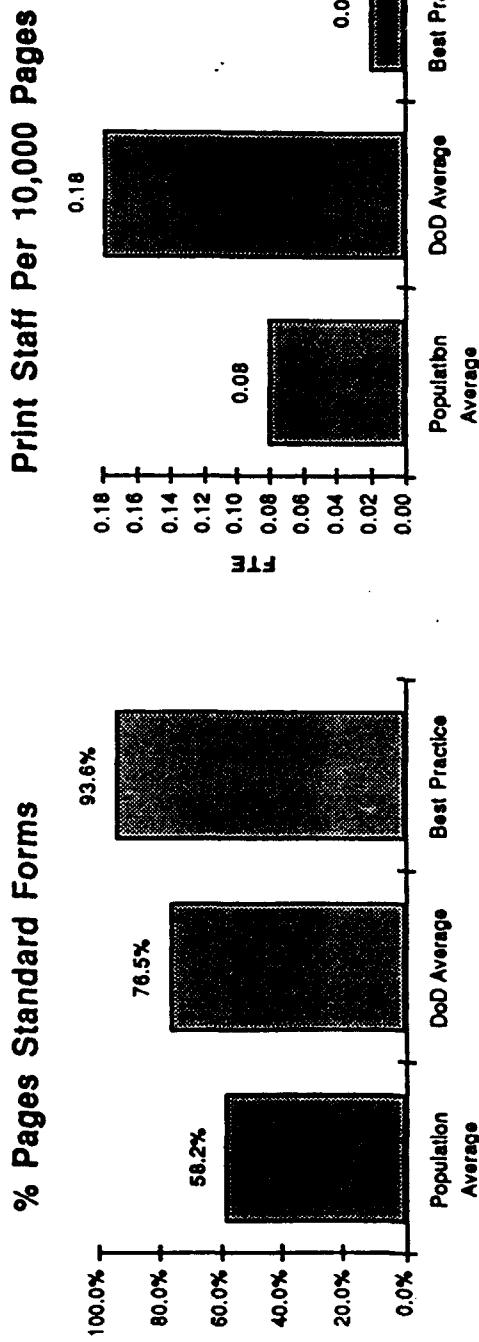
% Production Failures



©Copyright 1992 by KPMG Peat Marwick. All rights reserved.

KPMG Peat Marwick
Management Consultants

Figure 5: DoD DPI Abend Rate



The standard at DoD sites is one image per page.

©Copyright 1992 by KPMG Peat Marwick. All rights reserved.

KPMG Peat Marwick
Management Consultants

Figure 6: DoD Print Productivity

Control Software Products (Degree Of Implementation On A Scale Of 0 - 100)	DoD MVS Average	Database Average	Variance From Average	Best Practice
Console Management	27.3	41.2	-13.9	70.0
Tape Management	47.9	73.1	-25.2	90.0
DASD Management	56.6	50.2	+6.4	80.0
Security	61.0	69.2	-8.2	90.0
Scheduling	47.2	68.5	-21.3	90.0
Restart/Rerun Management	30.8	34.6	-3.8	90.0
Output Distribution	46.3	52.5	-6.2	82.0
Balancing	0.0	13.3	-13.3	72.6
Problem Management	25.2	53.1	-27.9	80.0
Change Management	19.6	48.1	-28.5	80.0
Configuration Management	14.2	26.9	-12.7	80.0
Software Management	15.0	58.7	-43.7	80.0
Data Center Reporting	34.0	49.0	-15.0	92.6

Figure 7: DoD DPI Operations Automation

Peat Marwick also found that DoD DPIs are far behind the population average when compared on the utilization of automation software. As can be seen in Figure 7, DoD DPIs are only above average in one of thirteen categories: DASD management. One of the primary reasons that DoD DPIs are less efficient than their counterparts in the Peat Marwick data base is that they employ more staff than the population average. As can be seen in Figure 8, DoD DPIs employ over 100% more staff per used MIPS than the population average, and over 300% more staff than the best practice. The result of this higher staffing is shown in Figure 9, higher cost.

The principle problem areas that appear to be the primary drivers for the above results are:

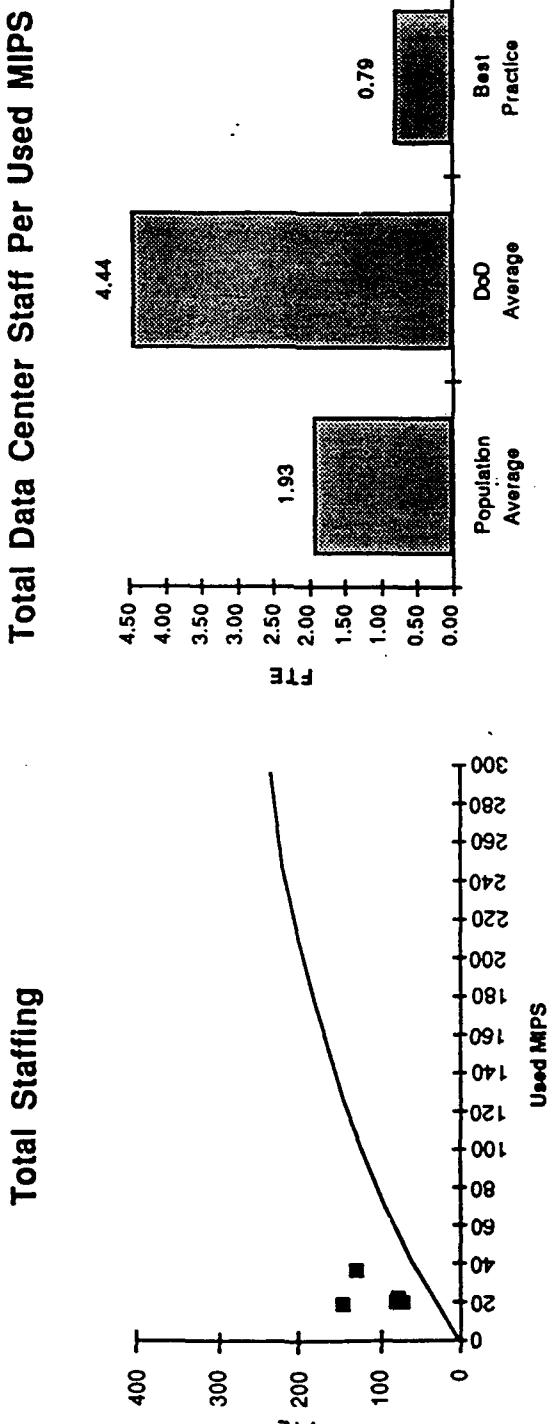
- DoD DPIs have multiple computer operating environments, each with its own unique operating system.
- Even within a single operating environment, DoD DPIs have multiple small computers instead a few large computers.
- DoD DPIs do not use automation software to minimize the amount of staff needed.
- DoD DPIs do not have control over the job control language of the applications that they run. This minimizes their ability to ensure that the applications use the DPIs efficiently.

In general, the Pilot Project has identified the drivers that will enable DoD DPIs to become as efficient as other world class DPIs. These drivers are (1) rigidly enforced standards that ensure that all applications use the DPIs efficiently and (2) heavily utilized automation software to minimize the amount of staff needed and also to ensure that applications comply with established standards.

3. Part 2: Benchmarking and DPI Economics

3.1 Measuring DPI Size

The hardest part of comparative analysis is establishing valid comparisons. Each organization is different, as all people are different. All organizations are similar, as all people are similar. The benchmark vendors establish a model of a data center. Special studies are conducted to populate the model. If all or part of an essential function, say technical support, is provided from a central facility, the staff and budget assigned to the DPI are collected from the central facility and added in. If the DPI does things not in the model, say applications programming, the staff and overhead are subtracted. The result may not be a comprehensive picture of the DPI, but it is a comparable picture of the activities being studied with other DPIs.



©Copyright 1992 by KPMG Peat Marwick. All rights reserved.

KPMG Peat Marwick
Management Consultants

Figure 8: DoD DPI Staffing Levels

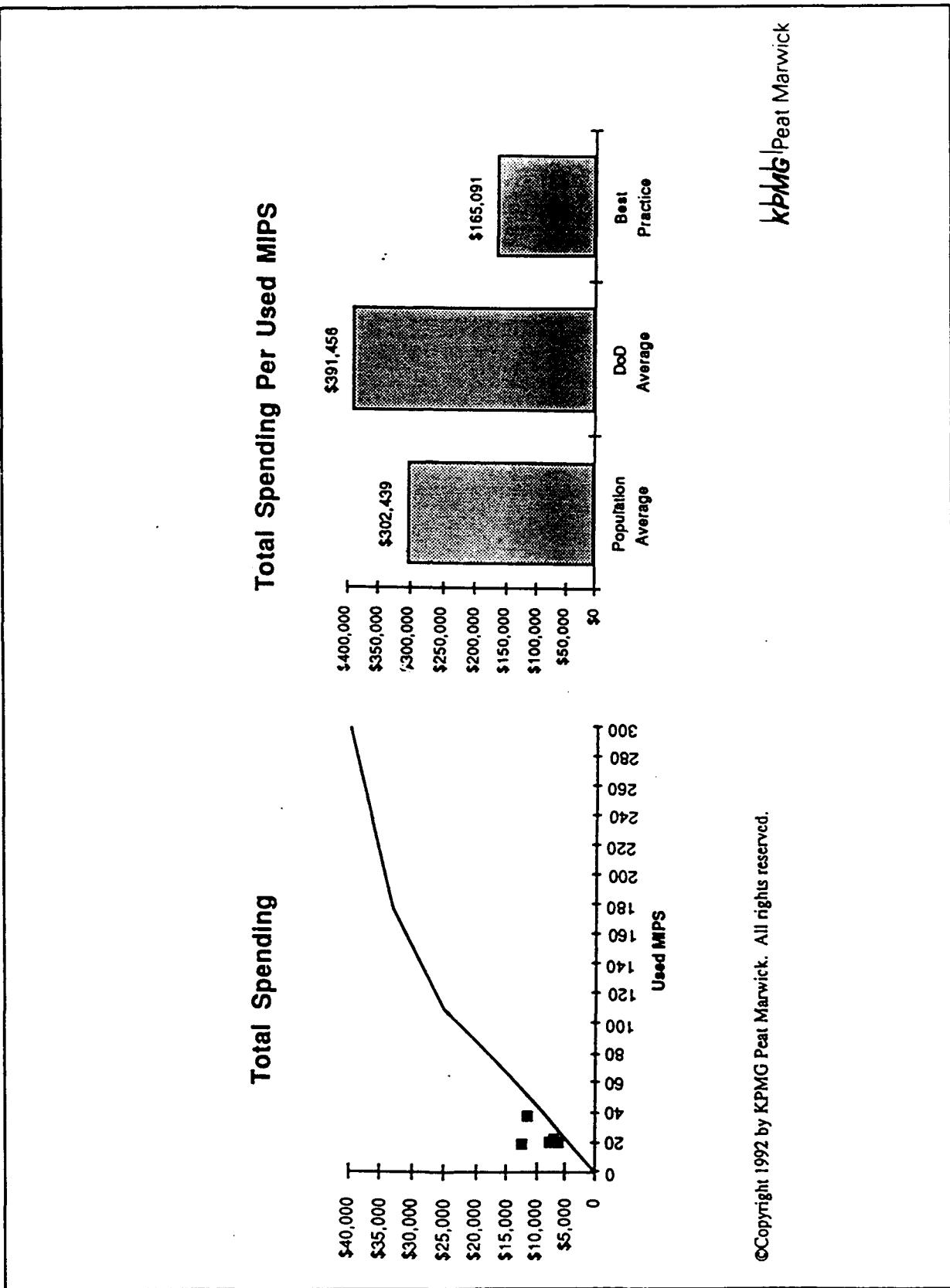


Figure 9: DoD Total Cost

Allowance must also be made for the workload. Big DPIs need more staff. The issue is how to measure the workload in a way that allows comparisons across DPIs processing different applications serving different clients in different industries. There have been many attempts to construct metrics of computing power. These resemble acquisition benchmarking, they calibrate the amount of a fixed workload processed by the platform. There is, however, no agreed upon metric. Each vendor and applications community prefers those workloads which focus on their salient features.

One measure of the capability of a DPI is the power of the Central Processing Units (CPU) measured in MIPS. The failings of MIPS are well known. Comparing DPIs on the basis of MIPS is rather like comparing motor vehicles on the displacement of their engines. In such a situation a race car may be considered equal to a truck. The CPU does not do the work alone. Input, output and storage devices are needed. The peripherals needed for a process control application differ from that of commercial processing. Furthermore, an instruction on one machine may not be of equal power as an instruction on another, just as a cubic inch of displacement produces different power on a turbo-charged engine.

These concerns can be handled the same as they are in motor vehicles. Vehicles of similar class have similar properties. Race cars can be compared on engine size alone. The same applies to DPIs. In particular, IBM-compatible DPIs doing commercial data processing using applications which run under the MVS operating system are sufficiently similar to permit comparison based on MIPS alone without paying close attention to the differences in their specific instruction set or suite of peripherals. Comparisons to PCs or Workstations or to non-IBM compatible machines require such extensive normalization as to make the effort of dubious value.

In sum, raw computing power sitting on the floor does not represent work accomplished. Raw computing power is also subject to swings as machines are upgraded. Used CPU power is more stable. For this reason the vendors use some variety of used processing capability for comparison. Real Decisions bases much of their analysis on billable MIPS. This is the amount of computer time used by customer applications net of overhead processing and idle time. Nolan, Norton converts the several elements of computing capability into their measure of computing power called a RIP. They then discard idle time to get Used RIPS (URIP). Compass sizes DPIs on installed MIPS, configured MIPS (a slight reduction from installed MIPS), and used MIPS. Compass also frequently normalizes based on the CPU seconds which would be consumed had the work been done on a theoretical 1 MIPS machine. KPMG reduces the stated MIPS value by CPU busy to obtain used MIPS.

The vendors use slightly different definitions of MIPS; MIPS itself is subject to considerable estimation error; the scope and detail of the functions varies between vendors. As a result, close comparison of their results are not possible. Nolan, Norton's use of the RIP avoids the debates regarding how to normalize MIPS, but raises equal questions regarding the validity of the RIP. Since the important factor is the DPI's position relative to

others in the same size class, it is of little consequence whether the analysis is based on RIPS or MIPS. The vendors take care to ensure comparability between their clients. However, there is no comparability between the vendors given their different perspectives.

3.2 Vendor Perspective

The results of a benchmark strongly depend on the weltanschauung of the vendor. Due to the differences in perspective it would not be improbable for two vendors to render conflicting judgments of a DPI.

The use of Third-Party benchmarking vendors has several advantages over attempting to tailor a benchmarking study to the DPI. First, the metrics are predetermined. The essential functions of a DPI have been examined and operational definitions developed. This what-should-we-look-at issue is a major obstacle when conducting a tailored benchmarking study. Secondly, the comparison is made to multiple sites. In a tailored study, data must be collected on each benchmark partner. As a practical matter, this leads to one-on-one analysis. The ability to compare against multiple DPIs also permits comparison against multiple standards. Commonly the DPIs metrics are arrayed against DPIs of the same MIPS size. Additional comparisons can be drawn against DPIs in the same industry, e.g. banking, to account for differences in culture or application. A "threat assessment" can be conducted by comparing against the largest DPIs, surrogates for outsourcing firms. Comparison can also be made against Best Practice regardless of industry or size. And last, the DPI's identity is protected by the vendor. No other entity knows what the DPI's metrics are. They are reported only as averages or under code names.

Each of the vendors report on the core metrics: total cost per MIPS/RIPS/CPU second, total headcount per MIPS/RIPS/CPU second, CPU Utilization, tape/DASD/print management. Beyond these basic statistics they differentiate themselves. Real Decisions places great emphasis on costing and rate recovery. Nolan, Norton conducts interviews to assess the level and strength of management controls. Compass collects many detailed statistics on DPI operations. KPMG's strong suite is guidance provided the DPI staff toward improved operations.

Each vendor's focus is distinct. Real Decisions takes a grand strategic view. They provide to the CEO and CFO information on resource allocation. Nolan, Norton takes a strategic look which includes consideration of morale and sense of mission. Compass works at the operational level where the capabilities of the several DPI functions must be coordinated. KPMG focuses on tactical improvements within the functions.

The purpose underlying the analysis varies by vendor. Real Decisions provides executive insight. Nolan, Norton takes a holistic view which examines tangible and intangible aspects. Compass stresses the need for process improvement. KPMG's emphasis is on proper use of technology to attain state of the art operations.

The approach taken by the vendors reflects these basic differences in focus and purpose. Real Decisions applies the tools of financial management. Nolan, Norton's unique contribution is use of management analysis techniques. Compass takes the approach of industrial engineering. KPMG concentrates on technology management.

The vendors clearly differ in style. Selection of a vendor should turn on the style of management favored by the DPI or the business question leading to the study. Too much should not, however, be made of these differences. The primary value of benchmarking -- perspective -- is provided by all.

3.3 Measuring Overall Performance

The vendors collect many statistics about DPIS and report comparative metrics on a variety of functions. Together these form a comprehensive picture of the DPI. Clients, especially top management, want a single measure of the DPI. This single measure or grade is easily briefed up the chain and facilitates year-to-year tracking. Real Decisions and Compass have metrics based on the expense of providing a standard level of processing. KPMG's approach resembles an academic grade. They score several DPI functions and assign points. The perfect score would be 100. Nolan Norton does not have an identifiable single metric. Total Cost per Used RIP meets this need.

Real Decisions runs a suite of workload benchmark routines and records the resources consumed. From this they construct a fee for the work under a cost-recovery chargeback. This imputed bill is divided by the fee which would be charged by the median, in terms of efficiency, DPI. By this scheme the index for the median DPI is set to 1.0. This index is called the NOW. Roughly speaking, a DPI with a NOW of 0.8 would do for \$8,000 a workload that would command an average price of \$10,000. Conversely a DPI with a NOW of 1.2 would consume resources worth \$12,000 to accomplish the work.

Compass's overall metric is called the HUGO. It is the cost of 1000 CPU seconds of peak adjusted, external load normalized to a one MIPS machine. Hence the HUGO costs the demand which most likely was crucial in sizing the facility.

Scalier economies can be examined by displays of Total Cost per Used MIPS, Total Cost per Used RIPS, Total Cost per CPU hour, NOW or HUGO as a function of DPI size. Total Cost per Used MIPS or Used RIPS are adequate substitutes for the NOW or the HUGO as summary statistics or for year-to-year tracking.

3.4 Industry Economics

The vendors use graphics to display the results. These can be examined to develop a feel for the economics of DPIS. The data confirms the generally accepted belief of economies of scale in data processing -- the unit cost decreases as the DPI gets bigger. A closer look at the data, however, suggests the prime cause of the savings is not in the

underlying technology, but rather in management of the technology. Size does help. Hardware and software prices show modest economies of scale for larger machines, but skill is the main factor.

KPMG casts a trend line through the data. The DPI is then plotted as a point above, below or on the trend line. This lets the DPI understand how it is performing relative to industry norms. The trend line shown in Figure 10 says the industry exhibits returns to scale. Figure 11 implies the source of these returns is in staffing. When Nolan, Norton divides the budget, Figure 12, between technology and personnel, there is a sharp large/small distinction. The larger centers spend proportionally less on people, despite a well known trend toward higher salaries in large centers. These data argue for, and explain the source of, economies of scale in DPI operations.

Trend lines follow the average. If average performance is acceptable, economies of scale are definitely present. If the objective is superior performance, scale recedes in importance. As noted in Figure 13, efficiency varies considerably across the industry. The best DPI is ten times as efficient as the worst. Among the top performers, however, there is little spread. Figure 14 displays the NOW index of Real Decisions Best Standard of Excellence (BSE) group. This is a group of DPIs who process a typical workload. (The DPI with the best NOW was excluded from the BSE due to the beneficial impact of an unusual workload.) Since some variation is to be expected, these DPIs are equally efficient and demonstrate the existence of a definable state of best practice. The DPIs in the BSE range in size from small to very large. Clearly scale is not the only factor at work.

Consider the display of HUGO by MIPS, Figure 15. HUGO tends to improve as MIPS increase, but note the wide variation of performance for close values of MIPS. Scale is not dominant. Most importantly, observe the DPIs in the most efficient position, the bottom of the chart. A bounding line placed at the bottom of the data would define the efficient frontier of operations. The efficient frontier does not show economies of scale. The curve to a trend line placed through the data comes, not from the best performers, but from the large number of poor performers. The number of inefficient DPIs decreases with DPI size. This implies a return to management, not scale.

Many factors contribute to the dominance of management in DPI efficiency. First, it is easy to operate an inefficient, small DPI. A large staff can accept ill-formed work from the clients and make the adjustments to complete the job. It is difficult to conduct large scale operations with gross inefficiency. A large DPI must insist on adherence to standards or it would simply be unable to process the volume of work. Second, good managers do not tolerate overstaffing. Third, good managers purchase hardware and software at advantageous prices and employ the correct technology, neither retaining obsolete equipment nor paying a premium for the most recent advance in the state of the art.

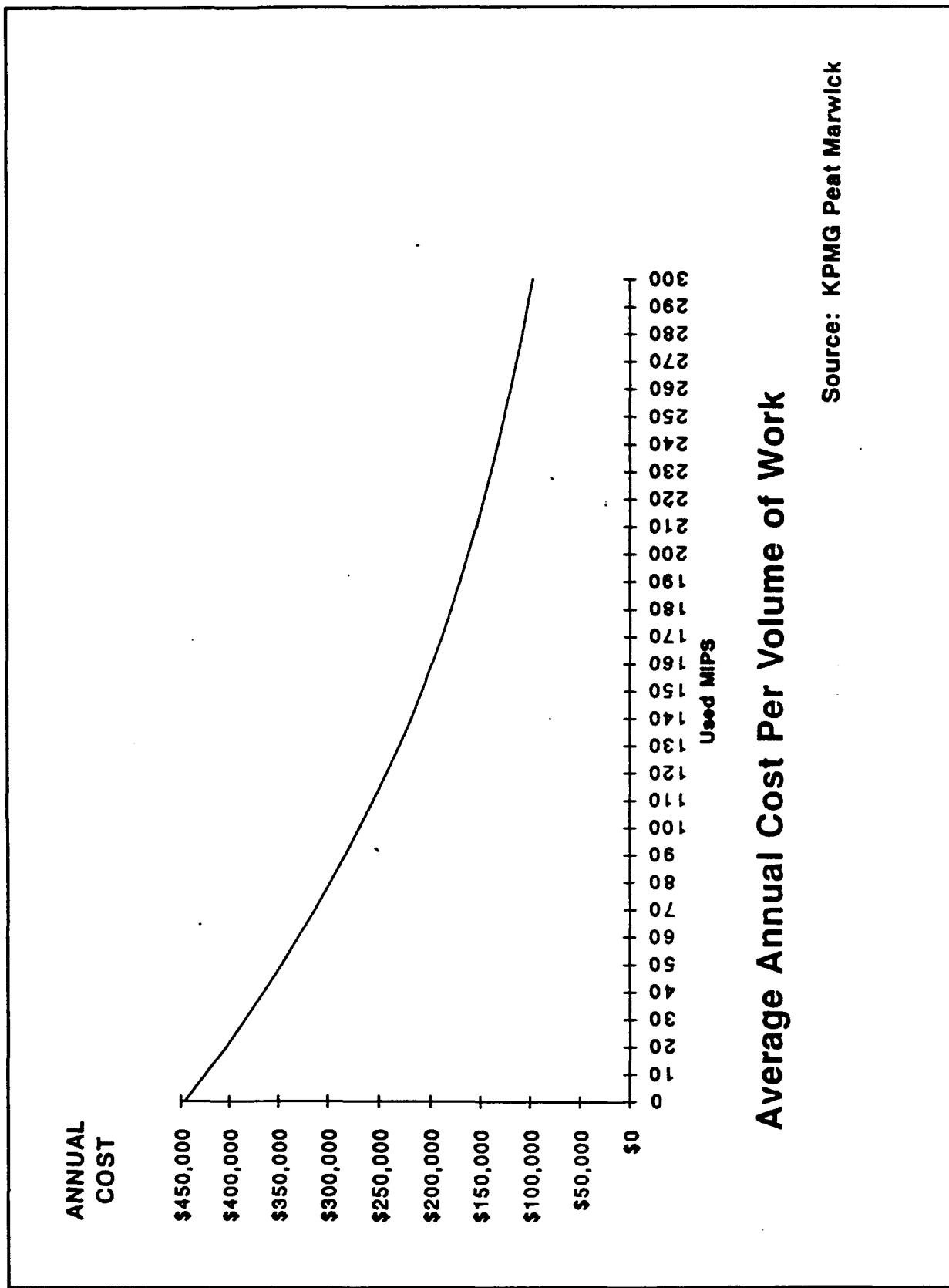


Figure 10: Industry Returns to Scale

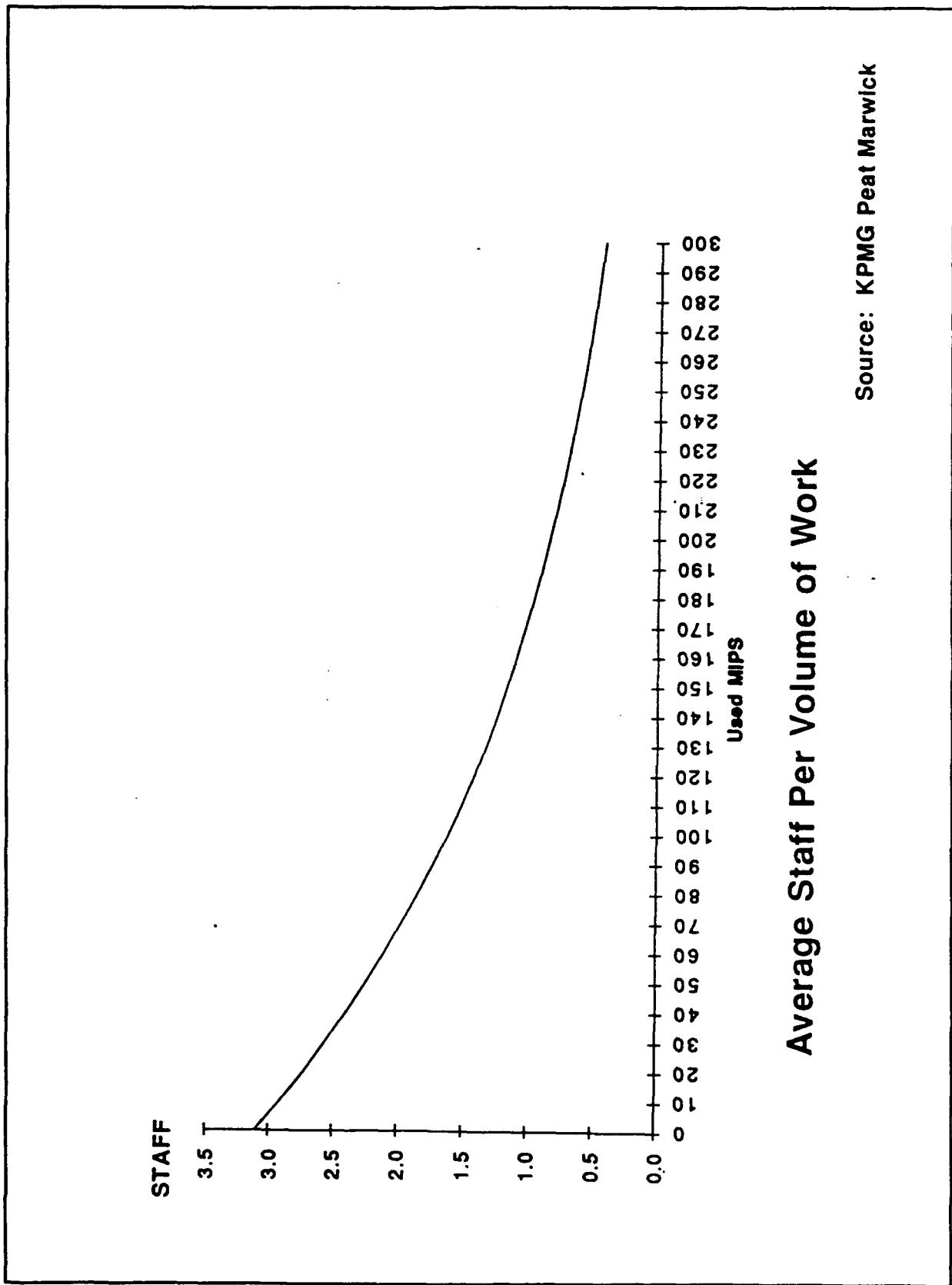
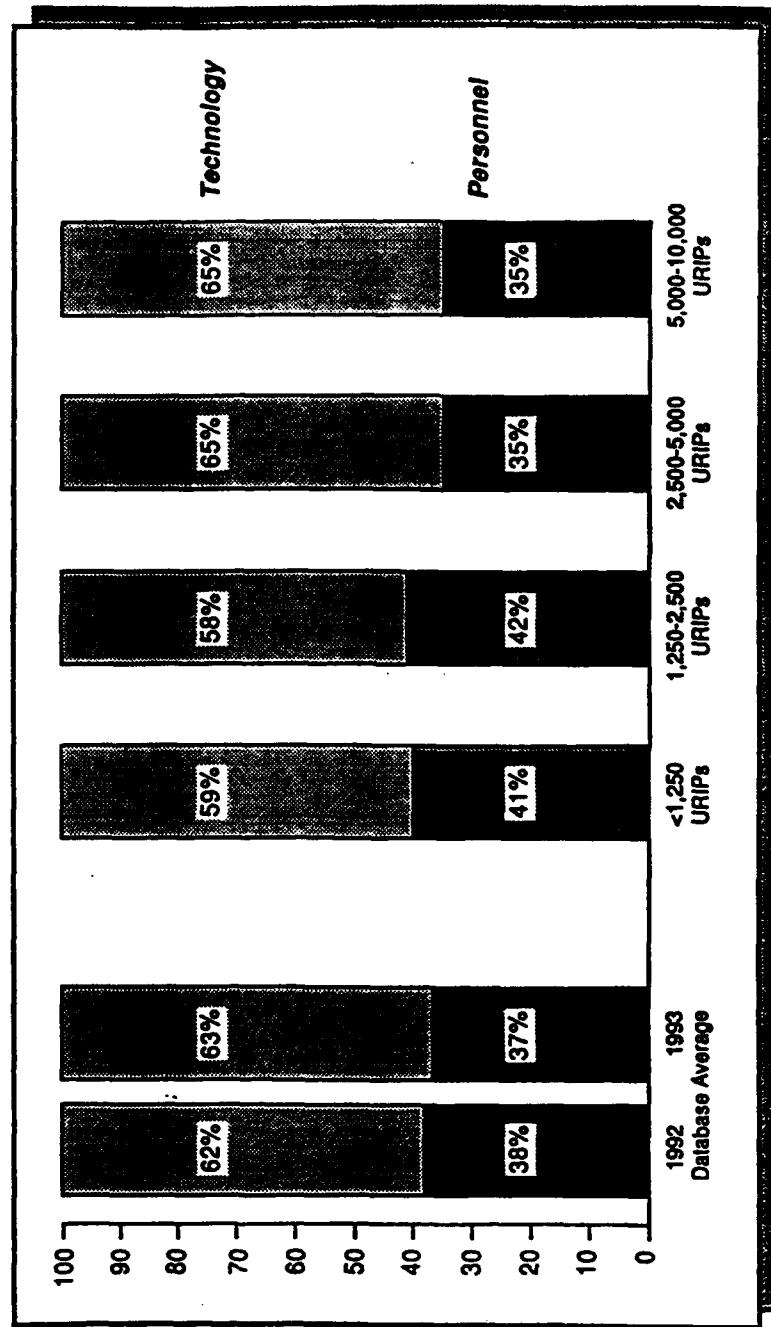


Figure 11: Staffing Levels Per Used MIPS



- Smaller data centers tend to spend more on personnel

©1993 by Nolan, Norton & Co. All rights reserved

Figure 12: Data Center Spending Mix

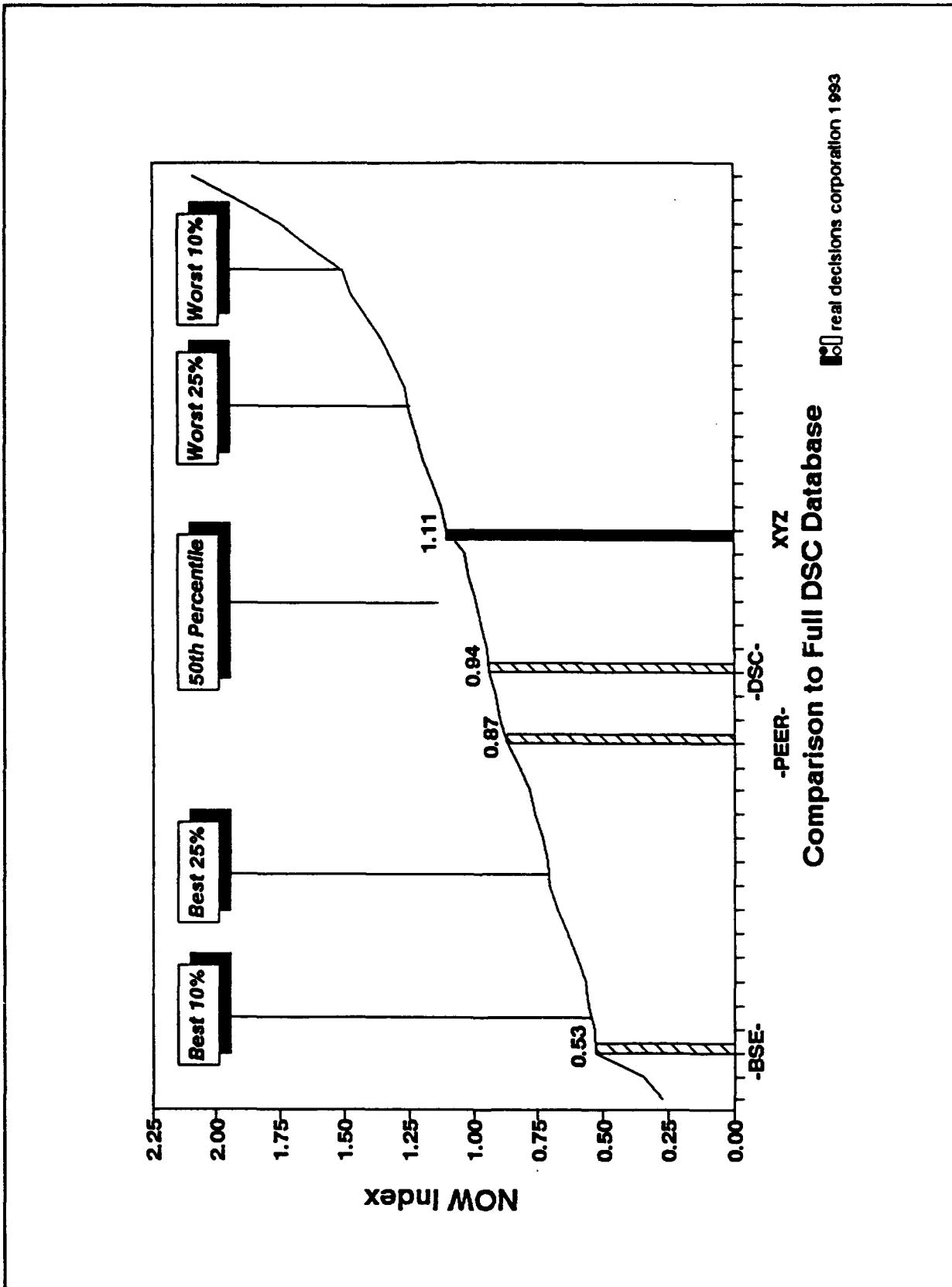


Figure 13: NOW Index Comparison

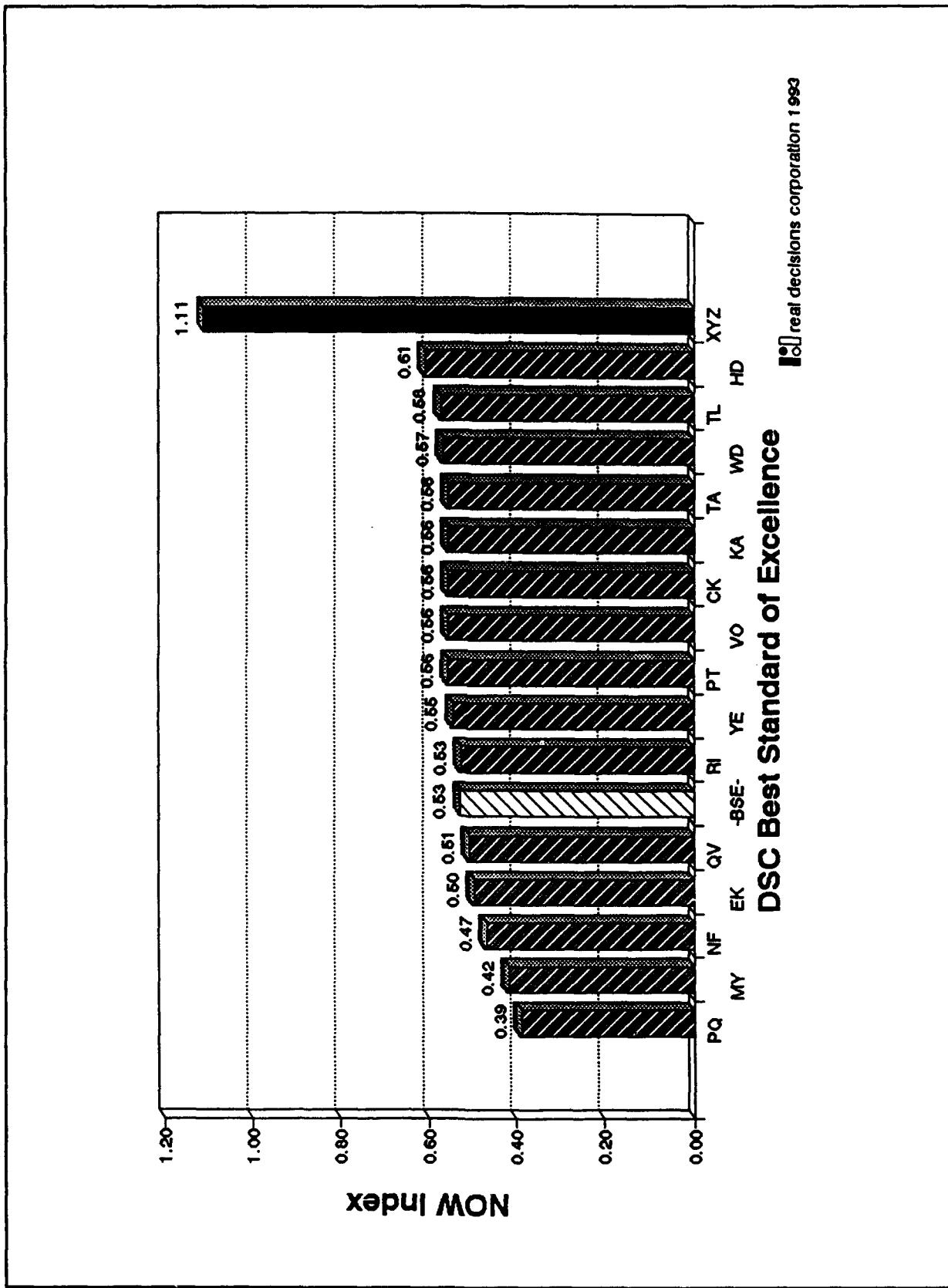
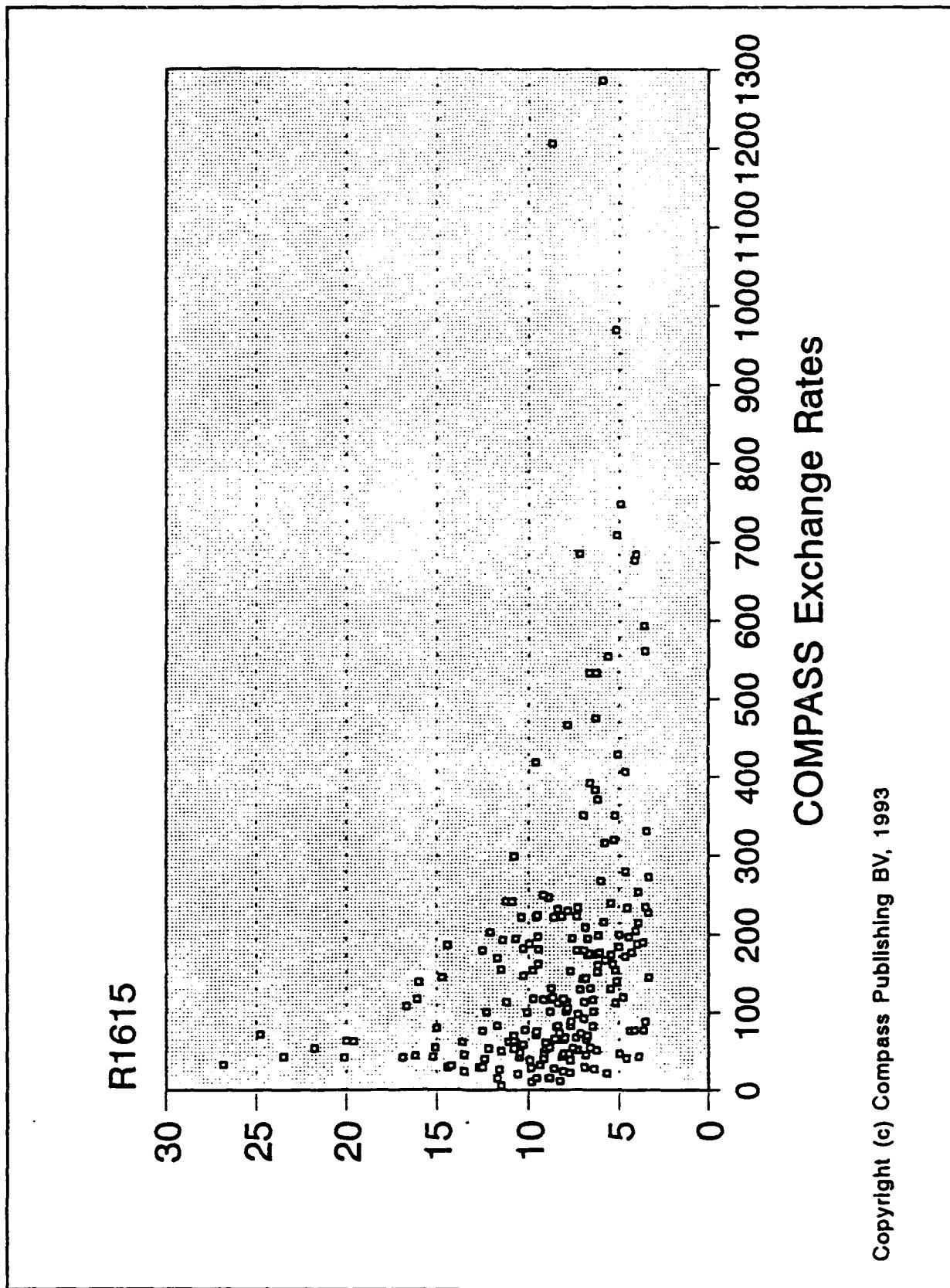


Figure 14: BSE NOW Index Distribution



Copyright (c) Compass Publishing BV, 1993

Figure 15: HUGO Distribution by MIPS

The latest topic in DPI management is rightsizing -- operating with staff, hardware and software which precisely meets the requirements of the processing load. There is no excess capacity in the overhead charge. Rightsizing takes considerable management skill. Since large DPIs pay higher salaries, they get the better managers and hence the existence of returns to scale. The dominant driver, however, for these scalar economies is not mere scale, it is managerial skill.

A bounding line placed at the bottom of the data indicates the potential for dis-economies of scale for very large centers. This may be the case, but it is too soon to tell. DPIs of this size have only recently begun to appear. They may drop to the efficient frontier once management learns how to control them.

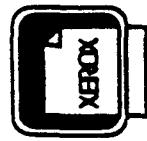
This analysis supports a policy of consolidating DPIs. Consolidation saves money by eliminating needless redundancy within the organization. There are economies of scale in the hardware and software in terms of the price paid after skillful negotiation with the vendors. Operating multiple centers entails coordination costs to control the configuration of enterprise-wide data and applications. There is some duplication of staff which can be avoided if all the work is done in one place. But the strongest argument for consolidation is that good managers are scarce. If you have lots of DPIs, you have lots of opportunities to have a bad manager in charge.

Management aside, the amount of reserve capacity needed for contingencies is smaller for one large center than the sum of such capacities for multiple small centers, i.e. a single large center is *more rightsized*. Some applications work best using technology like Automatic Tape Libraries which demand large volumes of work to justify. It is simply cheaper to obtain the service required by the organization from one DPI rather than from many. Thus consolidation of several small DPIs into one large DPI under the best available manager results in savings.

This critique of the underlying causes of scalar economies does contradict the notion of outsourcing in order to capture economies of scale by combining your work with that of others. Some economies of scale do exist, but the outsourcer's primary contribution will be savings which come from applying standards, discipline, and automation to the work -- savings which could be achieved in-house. The value added is management, not scale. When you add in the cost of managing the outsource contract, the loss of control, and the risk of over/underestimating demand, outsourcing loses its appeal.

3.5 Conclusion

The Xerox Corporation is a strong advocate of benchmarking. They attribute the survival of the corporation in the face of severe competition to TQM. Benchmarking is one of two pillars in their TQM program. Based on clues found by benchmarking, Xerox has achieved a 20% per year compounded annual reduction in their cost per CPU hour, Figure 16. Analysis of their volume and cost, Figure 17, is interesting. From it, one can



Xerox Computer Center

Benchmarking: Performance 1987 - 1991

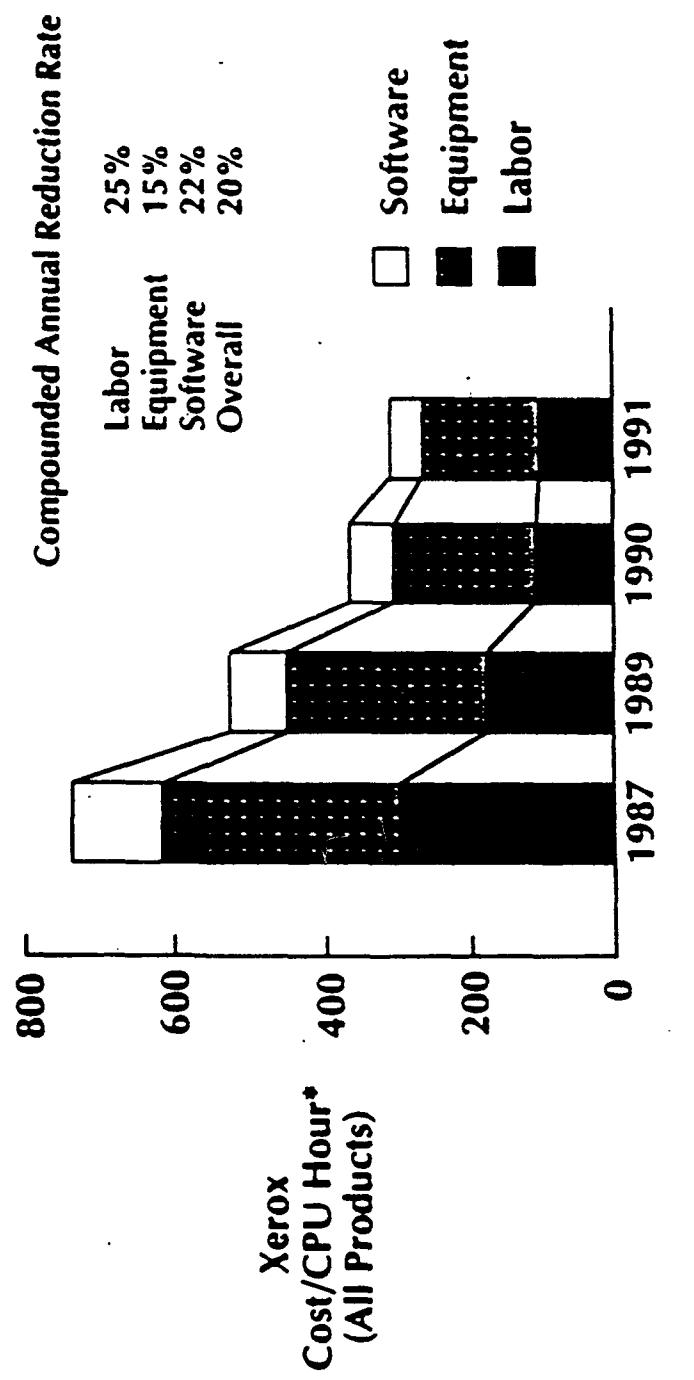
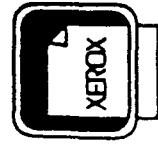


Figure 16: Xerox Benchmarking Performance



Xerox Computer Center

Results: Product Volume/Cost/Price Index

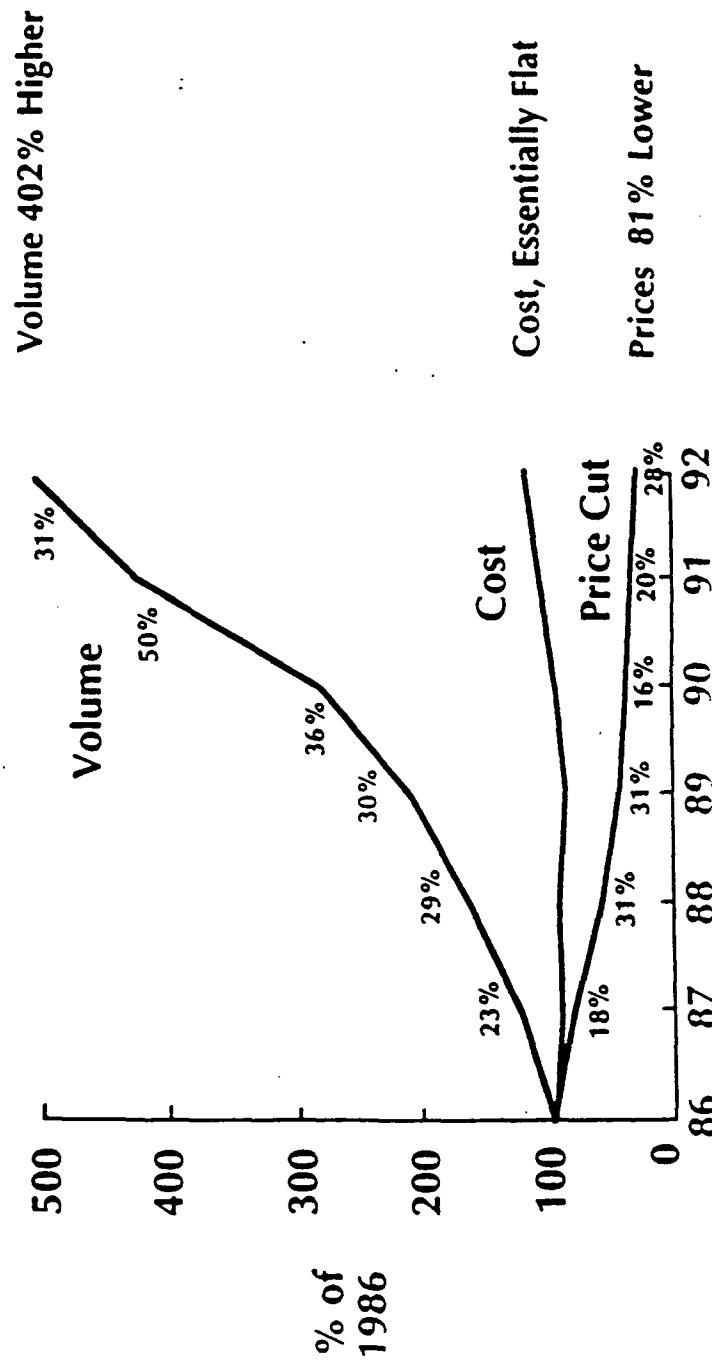


Figure 17: Xerox Volume, Cost, and Price Analysis